DOCUMENT RESUME

ED 262 097 TM 850 580

AUTHOR Goodison, Jules
TITLE National Assessment of Educational Progress: An Update of the Data Collection Process.
PUB DATE Apr 85
NOTE 17p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Chicago, IL, April 1-3, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; *Data Collection; *Educational Assessment; Educational Testing; Educational Trends; Elementary Secondary Education; *National Surveys; *Sampling; School Statistics; Student Evaluation
IDENTIFIERS *Balanced Incomplete Block Spiralling; *National Assessment of Educational Progress

ABSTRACT

Since 1969 the National Assessment of Educational Progress (NAEP) has conducted annual assessments to determine, in the language of the current legislation, "the performance of children and young adults in the basic skills of reading, mathematics and communications" and to "report data periodically on changes in the knowledge and skills of such students over a period of time." Since 1983, when Educational Testing Service (ETS) received the grant to administer NAEP, some important changes have been made in the original NAEP design to increase its utility, including: (1) sampling by grade as well as age, thereby allowing schools and states to compare NAEP data with their own grade samples; and (2) Balanced Incomplete Block (BIP) spiralling. Sampling became more efficient in BIP spiralling, since particular blocks of exercises were given to fewer students within a school, but were administered in more schools. Computation of correlation coefficients between all exercises was made possible in that each pair of blocks received a response. Individual items were still reported, and correlations were estimated from comparatively small samples. Diagrams of the spiralling process and tables depicting school cooperation and student participation accompany the text. (LMO)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED262097

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.
"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Goodison, J.
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS: AN UPDATE OF THE DATA COLLECTION PROCESS

Jules Goodison
Educational Testing Service

I. Introduction

As was indicated in presentations made at AERA/NCME last year in New Orleans, the National Assessment of Educational Progress (NAEP) has conducted annual assessments from 1969 through 1982 and biennial assessments from then until the present. The primary purpose, as stated in the current legislation, is to assess "the performance of children and young adults in the basic skills of reading, mathematics and communication" and to "report data periodically on changes in the knowledge and skills of such students over a period of time" (General Education Provision Act). Since 1969, over 1,100,000 youths have been assessed. The samples have consisted of 9-, 13-, and 17-year-olds who were in school. As funding would allow, the sample also included 17-year-olds who had left school and, on occasion, adults in the 26-35-year-old range. In 1983, following some 14 years of implementation by the Education Commission of the States (ECS), Educational Testing Service (ETS) received the grant to administer NAEP for the next five years.

ETS made some important modifications to the original NAEP design to increase its utility. Some of the most important changes were:

- Sampling Grade as well as Age

Historically, the NAEP assessed ages, not grades. By sampling ages alone, NAEP was unable to make important comparisons with other educational databases. Schools are organized by grade, and state assessments are by grade, and thus a school or a state could not safely use the NAEP data for performance comparisons. Collecting national samples for both grades and ages necessitated sampling about 30% more students than sampling by age only.

TM 850 580

- BIB Spiralling

Perhaps the most innovative part of the new design was the introduction of Balanced Incomplete Block (BIB) spiralling. An important NAEP objective has been to have a broad coverage of each area that is assessed. In the past, NAEP has used matrix sampling; that is, the total assessment time was divided into mutually exclusive booklets of exercises, each of which contained about three-quarters of an hour of exercises. This basic matrix design allowed for computing correlations and cross-tabulations among exercises within the same booklet, but did not allow for correlations and cross-tabulations among exercises within different booklets. The 1983-84 ETS design, however, divided the total assessment battery into blocks of 14 minutes each -- 2 minutes of background questions and 12 minutes of cognitive items. Each student was assigned three blocks, as well as a six-minute block common to all students, so that the total assessment time for each student still approximated three-quarters of an hour. What BIB spiralling does is to generate a very large number of booklets, but does so in such a way that each block of exercises containing both background and cognitive questions is paired with another block in one, and only one, booklet. These booklets are then spiralled so that few students in any one assessment session are given the same booklet.

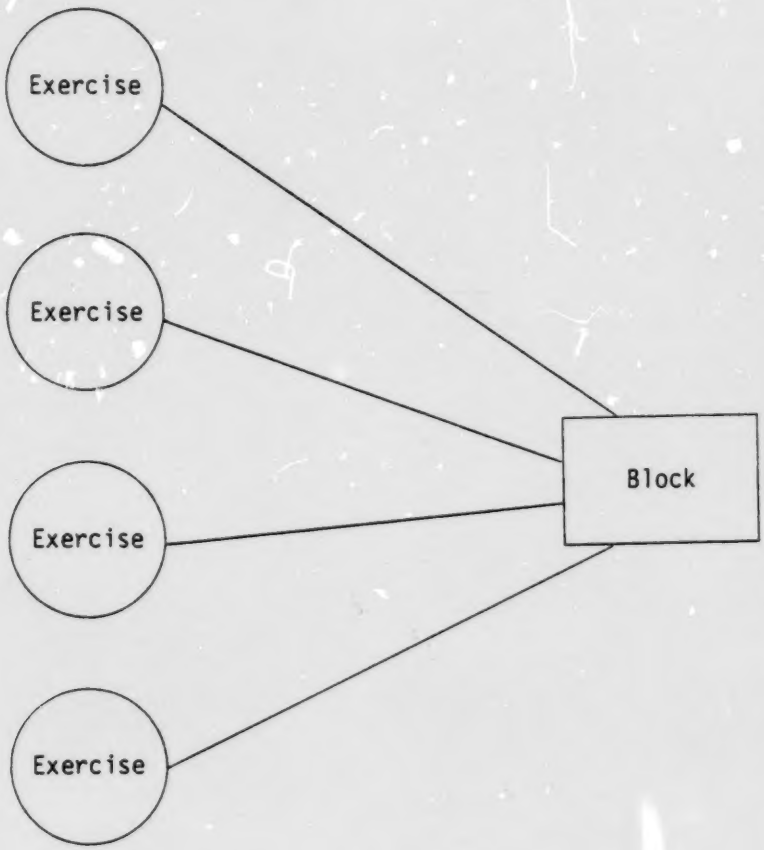
The major advantage of BIB spiralling is that it allowed for the computation of correlation coefficients between all exercises, since we are guaranteed that some students have responded to every pair of blocks, and still allows us to report on individual items as was done in the past. It is worth noting that most correlations are estimated from comparatively small samples. BIB spiralling makes it possible to attempt to calibrate all exercises in the assessment using item response theory. BIB spiralling also makes sampling more efficient, since particular blocks of exercises are given to fewer students within a school, but administered in more schools.

II. The Spiralling Process

NAEP divided the item pool into twenty-four blocks of exercises. Twenty-one single-length blocks each with 14 minutes of testing and three double-length blocks each with 28 minutes of testing. The design for the 1983-84 Assessment-Balanced Incomplete Block (BIB) and Unbalanced Incomplete Block (UBIB) assured that each possible pair of blocks occurred in at least one book so that inter-block analyses would be possible. One of the other design factors of the new NAEP was to collect bridge samples which administered some of the NAEP exercises using audio recordings and simple matrix sampling, as was done in past assessments. The purpose of these samples was to explore the effect of the change from audio recordings to pencil-and-paper administration and, if possible, to project the results of past NAEPs onto the new scale.

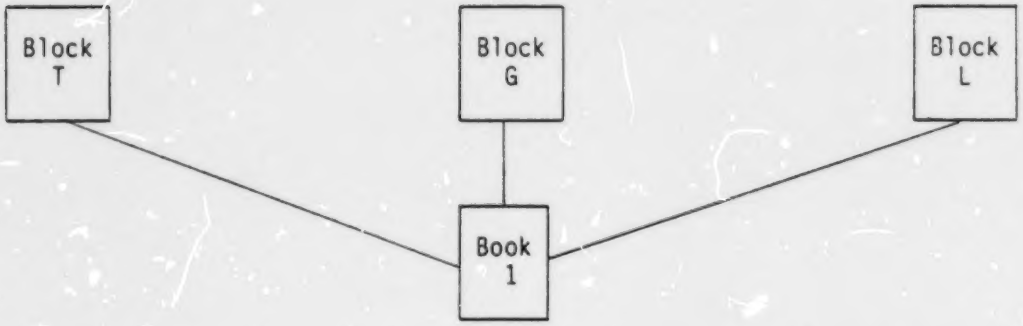
The spiralling process can be diagrammed as shown on the next page:

Figure I



Then 3 blocks of exercises are assembled into an assessment book as diagrammed in Figure II.

Figure II



The first nineteen of the single-length blocks were assigned to books in threes. This produced 57 distinct books, each possible pairing appears in one of the books, and each block appears in nine books. The block assignments are shown in Table I.

Table I

NAEP Book Design
BIB Spiral Sample

Book	Item Block 1	2	3	Book	Item Block 1	2	3
1	T	G	L	29	Q	O	D
2	A	L	P	30	B	Q	J
3	D	A	T	31	O	T	H
4	C	S	E	32	B	M	L
5	C	A	H	33	C	R	O
6	G	F	H	34	G	O	E
7	K	R	N	35	S	Q	M
8	R	M	F	36	B	A	O
9	O	N	L	37	K	G	A
10	F	D	B	38	O	F	K
11	E	M	A	39	P	S	T
12	S	H	B	40	E	F	L
13	M	K	D	41	H	M	J
14	T	N	J	42	J	E	D
15	M	T	C	43	F	J	A
16	C	L	Q	44	B	G	C
17	H	E	R	45	P	B	K
18	C	P	F	46	S	F	N
19	L	S	K	47	P	Q	E
20	N	B	E	48	B	R	T
21	N	C	D	49	P	M	O
22	Q	K	H	50	R	P	D
23	L	H	D	51	G	R	Q
24	A	S	R	52	S	G	D
25	L	J	R	53	H	P	N
26	T	F	Q	54	T	E	K
27	C	K	J	55	M	G	N
28	O	J	S	56	A	N	Q
				57	G	J	P

Six more books were constructed to accommodate the remaining item blocks. Two of the BIB blocks, J and R, the two remaining short blocks, Y and X, and the three double-length blocks, U, V, and W were used. The two BIB blocks were added to the UBIB design to provide a link between the two item block types. In grade 4/age 9, block Q was substituted for block Y. The finished design for the UBIB is shown in Table II.

Each of the four Pace books have three sections, the common background section, a section of cognitive items, and a section of background and attitude items.

Table II

NAEP Booklet Design
UBIB Sample

Book	Long Block	Short Block
58	U	J
59	U	Y
60	V	R
61	V	X
62	W	X
63	W	Y

The spiralling of the NAEP books was designed to accomplish two purposes:

1. To achieve a distribution of nine students taking each of six UBIB books and two students taking each of 57 BIB books;
2. To distribute all books evenly across the sample of students so that no one book would be favored over another, except as described above under 1.

The first purpose was accomplished by forming a cycle of 168 books consisting of two sets of BIB books (1-57) and nine sets of UBIB books (58-63). The BIB and UBIB books were merged as follows:

1 2 58 3 4 59 5 6 60 7 8 61 9 10 62 11 12 63 13 14 58 15 16
59 17 18 19 60 20 21 61 22 23 62 24 25 63 26 27 58 28 29 59 30 31 60
32 33 61 34 35 62 36 37 38 63 39 40 58 41 42 59 43 44 60 45 46 61 47
48 62 49 50 63 51 52 58 53 54 59 55 56 57 60 1 2 61 3 4 62 5 6
63 7 8 58 9 10 59 11 12 60 13 14 61 15 16 62 17 18 19 63 20 21 58
22 23 59 24 25 60 26 27 61 28 29 62 30 31 63 32 33 58 34 35 59 36 37
38 60 39 40 61 41 42 62 43 44 63 45 46 58 47 48 59 49 50 60 51 52 61
53 54 62 55 56 57 63

Any given BIB book occurs twice in this cycle; any given UBIB book, occurs nine times. As long as this cycle of books is administered evenly across the sample of students, the desired ratio of nine UBIB books to two BIB books will be achieved.

The second purpose was accomplished by collecting this cycle of 168 books into bundles of 23 consecutive books, with consecutive bundles beginning where the previous bundle left off; the last of the 168 books is always followed by the first, in a continuous circling process (hence the term "spiralling"). The result was the creation of 168 different bundles with each book distributed evenly throughout the 23 positions in these bundles. By shipping consecutive bundles, the likelihood that any given book would be used was equalized across the sample.

The actual configuration of the 23 books in package #1 is as follows:

1 2 **58** 3 4 **59** 5 6 **60** 7 8 **61** 9 10 **62** 11 12 **63** 13 14 **58** 15 16

III. School Cooperation

The cooperation rate for schools participating with NAEP in the 1983-84 assessment year was outstanding. The original sample of schools for year 15 included 1,682 schools. District-level refusals to cooperate eliminated 142 schools. Another 53 individual schools refused to cooperate, 31 schools were closed, and 9 schools had had a change in grade span that made them ineligible for the grade/age class for which they were selected. Two other schools should not have been included in the sample because they had relocated and were no longer in the sample geographic areas. The remaining 1,445 agreed to participate and assessment sessions were conducted in 1,342. The difference between the number agreeing to participate (1,445) and the number in which sessions were held (1,342) is attributable to those schools which had no eligible students (98) and those where no eligible students were sampled (5). Schools included in the sample were estimated to have at least one eligible student on the basis of available grade span and enrollment information. Since the sample was selected prior to the beginning of the school year, some of the sample schools were found not to have any eligibles, or when the

prescribed sampling procedures were followed, no eligibles were sampled. Additional schools beyond those originally selected had to be added to the sample. The disposition of all schools in the sample for the three grade/age classes, is presented in Table III. The cooperating schools for both the original and total sample are presented in Table IV.

Table III

Disposition of Schools in NAEP Year 15 Sample

School Disposition	Grade/Age I (9 yr. old/ 4th grade)	Grade/Age II (13 yr. old/ 8th grade)	Grade/Age III (17 yr. old/ 11th grade)	Overall
Initially selected schools	700	588	394	1,682
Schools added	19	3	1	23
Total original sample	719	591	395	1,705
Assessments conducted	607	455	299	1,361
Schools lost by district refusal	61	42	40	143
School refusals	19	14	21	54
No eligible students	17	64	17	98
Closed or out-of-range	15	12	17	44
No eligible students sampled	0	4	1	5
Replacements for refusals	67	28	34	129
Assessments conducted	54	23	27	104
Refusals	5	2	6	13
No eligible students	5	3	1	9
Out-of-range or closed	3	0	0	3
Overall sample				
Assessments conducted	661	478	326	1,465
Refusals	85	58	67	210
No eligibles enrolled or sampled	22	71	19	112

Table IV

Number of Cooperating Schools in NAEP Samples

	Original Sample Year 15 1983-84	Total Sample Year 15 1983-84
Grade/Age Class I		
Assessments conducted	607	661
No age eligibles	17	22
No eligibles sampled	-	-
Total	624	683
Grade/Age Class II		
Assessments conducted	455	478
No age eligibles	64	67
No eligibles sampled	4	4
Total	523	549
Grade/Age Class III		
Assessments conducted	299	326
No age eligibles	17	18
No eligibles sampled	1	1
Total	317	345
All Classes		
Assessments conducted	1,361	1,465
No age eligibles	98	107
No eligibles sampled	5	5
Total	1,464	1,577

Overall, the number of schools cooperating was greater in Year 15 than in Year 13 (1,577 compared with 1,473) as was the number where assessments were conducted (1,465 and 1,349, respectively). By Grade/Age Class, in Year 15, assessments were conducted in 248 more Class I schools than in Year 13 but in fewer Class II and III schools (by 62 schools and 70 schools, respectively). These differences are due to design specifications. In Year 15, the goal was to have approximately 30,000 students assessed per grade/age class. Since high schools tend to be large and primary schools small, a larger sample of primary schools had to be included to yield equivalent numbers of students assessed. The cooperation rate for the Year 15 original sample was 88.1. Table V presents the comparison of Year 13 and 15 cooperation rate.

Table V

Number and Percent of Schools in Original Samples Cooperating in NAEP

Grade/Age Class	Year 13	Year 15
I (4th Grade/9 Year Olds)	395 (88.0)	624 (88.6)
II (8th Grade/13 Year Olds)	586 (89.2)	523 (90.3)
III (11th Grade/17 Year Olds)	409 (86.5)	317 (83.9)
OVERALL	1,390 (88.0)	1,464 (88.1)

Note: The cooperation rate is calculated as follows:

$$R = \frac{\text{Assessments conducted + No eligibles}}{\text{Assessments conducted + No eligibles + Refusals}}$$

IV. Student Participation

The student participation rate in the 1983-84 NAEP assessment exceeded expectation. Of the 122,589 students sampled, 8,058 or 6.6 percent were ineligible or excluded by the school as educable mentally retarded, limited English-speaking ability or functionally disabled. Table VI presents the number of students sampled, the number ineligible or excluded, and the number to be assessed. Table VII presents the number and percent of sampled students who were assessed.

Table VI

Number of Students Sampled, Excluded, or Ineligible, and To Be Assessed in Year 15

	Grade/Age Class			
	4th grade/ 9 years old	8th grade/ 13 years old	11th grade/ 17 years old	Overall
Spiral Sessions				
Sampled	30,289	33,487	38,044	101,820
Ineligible or Excluded	2,019	1,868	3,060	6,947
To be assessed	28,270	31,619	34,984	94,873
Tape Sessions				
Sampled	6,325	6,147	8,297	20,769
Ineligible or Excluded	341	239	531	1,111
To be assessed	5,984	5,908	7,766	19,658
Both Types				
Sampled	36,614	39,634	46,341	122,589
Ineligible or Excluded	2,360	2,107	3,591	8,058
To be assessed	34,254	37,527	42,750	114,531

Table VII

Number and Percent of Sampled Students Who Were Assessed

Grade/Age Class	To be Assessed Year 15	Assessed Year 15
4th Graders/9 Year Olds		
Numbers	34,254	31,686
Percent	100.0	92.5
8th Graders/13 Year Olds		
Number	37,527	33,886
Percent	100.0	90.3
11th Graders/17 Year Olds		
Number	42,750	35,155
Percent	100.0	82.2
All Grades/Ages		
Number	114,531	100,727
Percent	100.0	87.9

Note: Year 15 data are preliminary.

V. Results of BIB Spiralling

Did "BIB Spiralling" work? Administratively the answer is an unqualified yes. As can be seen in Table VIII, Distribution of Booklets, the actual number of booklets used turned out to be evenly distributed. For books 1-57, we anticipated 289 books being completed. The actual number of completed books ranged from 290-322 for grade 4/age 9, 314-351 for grade 8/age 13, and 327-363 for grade 11/age 17. For books 58-63, with the 28 minute and 14 minute block, we anticipated 1,300 books being used. For grade 4/age 9 the range of books used was 1,385-1422, for grade 8/age 13 the range was 1,502-1,529, and for grade 11/age 17 the range was 1,528-1,572.

For the taped administration books where we anticipated 1,200 books being used, the actual usage for each grade/age was as follows:

Grade 4/Age 9	1,344-1,403
Grade 8/Age 13	1,276-1,310
Grade 11/Age 17	1,534-1,596

Table IX presents the distribution of blocks across all age/grade levels.

Blocks A-T were the 14-minute blocks and Blocks U-Y were the 28-minute blocks. There was no Y block for grade 4/age 9. Any greater usage for any one block was planned. Again, the answer to the question, did BIB spiralling work procedurally and administratively, the answer is an unqualified yes. Did BIB spiralling work psychometrically, and was it worth the administrative cost? My colleagues will answer these questions with an affirmative response.

Table VIII

Number of Books Administered
BIB and UBIB Spiral Sample

Book Number	Grade 4/ Age 9	Grade 8/ Age 13	Grade 11/ Age 17	Book Number	Grade 4/ Age 9	Grade 8/ Age 13	Grade 11/ Age 17
1	310	355	346	34	296	328	344
2	311	346	363	35	302	335	340
3	316	343	355	36	311	336	344
4	319	340	354	37	303	336	350
5	309	346	339	38	306	342	345
6	320	331	341	39	306	346	335
7	317	335	335	40	291	352	343
8	306	342	333	41	308	346	348
9	315	338	336	42	290	345	359
10	308	344	342	43	292	344	353
11	309	343	327	44	305	342	349
12	309	349	337	45	317	352	345
13	309	353	324	46	313	344	343
14	306	344	340	47	314	332	338
15	309	343	333	48	310	330	341
16	305	339	337	49	317	325	348
17	308	336	340	50	309	338	351
18	296	348	338	51	315	323	349
19	308	357	340	52	322	338	350
20	302	343	340	53	316	349	356
21	312	336	340	54	312	350	344
22	313	337	347	55	316	338	347
23	305	328	354	56	313	349	329
24	307	332	338	57	317	339	346
25	317	328	336	58	1,422	1,516	1,572
26	312	325	344	59	1,396	1,527	1,537
27	315	327	350	60	1,416	1,529	1,574
28	329	322	347	61	1,395	1,513	1,528
29	319	328	349	62	1,385	1,520	1,528
30	317	314	350	63	1,405	1,502	1,543
31	307	324	338	64	1,403	1,310	1,539
32	316	332	345	65	1,356	1,276	1,540
33	306	331	344	66	1,389	1,283	1,596
				67	1,344	1,289	1,534
				TOTAL	31,579	33,563	35,070

Table IX

Number of Blocks Administered
BIB and UBIB Spiral Sample

Block	Grade 4/ Age 9	Grade 8/ Age 13	Grade 11/ Age 17
A	2,771	3,075	3,098
B	2,795	3,042	3,093
C	2,776	3,052	3,084
D	2,790	3,053	3,124
E	2,741	3,069	3,089
F	2,744	3,072	3,082
G	2,804	3,030	3,122
H	2,795	3,046	3,100
J	4,213	4,525	4,701
K	2,800	3,089	3,080
L	2,778	3,075	3,100
M	2,792	3,057	3,045
N	2,810	3,076	3,066
O	2,806	2,974	3,095
P	2,803	3,075	3,120
Q	5,611	2,982	3,083
R	4,211	4,524	4,641
S	2,815	3,063	3,084
T	2,788	3,060	3,076
U	2,818	3,043	3,109
V	2,811	3,042	3,102
W	2,790	3,022	3,071
X	2,780	3,033	3,056
Y	0	3,029	3,080

References

Baratz, J.C. (1984). The New National Assessment of Educational Progress. Paper presented at the American Educational Research Association 1984 Annual Meeting, New Orleans, Louisiana.

Callahan, J. (Ed.) (1985) NAEP 1983-84 Public Use Data Tapes User's Guide Version 3.0. Educational Testing Service, Princeton, N.J.

General Education Provision Act, Title IV of PL 90-247, as amended by PL 95-561, Section 405(K)(1).

Westat, Inc. (1984) Preliminary Report on Field Operations and Data Collection Activities, NAEP - Year 15. Rockville, M.D.